

## Virtual Reality for Fire Evacuation Research

Max Kinateder  
Department of Cognitive,  
Linguistic, and Psychological  
Sciences  
Brown University, Box 1821  
Providence, RI 02912, USA  
Email: max\_kinateder@brown.edu

Enrico Ronchi & Daniel  
Nilsson  
Department of Fire Safety  
Engineering and Systems Safety  
Lund University  
P.O. Box 118, 22100 Lund,  
Sweden  
Email: enrico.ronchi@brand.lth.se,  
daniel.nilsson@brand.lth.se

Margrethe Kobes  
Institute for Safety  
PO Box 7010  
6801 AH Arnhem  
The Netherlands  
Email: Margrethe.Kobes@ifv.nl

Mathias Müller & Paul Pauli  
Department of Psychology I  
University of Würzburg  
Marcusstr. 9-11, D-97072 Würzburg, Germany  
Email: m.mueller@psychologie.uni-wuerzburg.de,  
pauli@psychologie.uni-wuerzburg.de

Andreas Mühlberger  
Department of Psychology, Clinical Psychology, and  
Psychotherapy  
University of Regensburg  
Universitätsstr. 31, D-93053 Regensburg, Germany  
Email: Andreas.Muehlberger@psychologie.uni-  
regensburg.de

**Abstract**—Virtual reality (VR) has become a popular approach to study human behavior in fire. The present position paper analyses Strengths, Weaknesses, Opportunities, and Threats (SWOT) of VR as a research tool for human behavior in fire. Virtual environments provide a maximum of experimental control, are easy to replicate, have relatively high ecological validity, and allow safe study of occupant behavior in scenarios that otherwise would be too dangerous. Lower ecological validity compared to field studies, ergonomic aspects, and technical limitations are the main weaknesses of the method. Increasingly realistic simulations and other technological advances provide new opportunities for this relatively young method. In this position paper, we argue that VR is a promising complementary laboratory tool in the quest to understand human behavior in fire and to improve fire safety.

### I. INTRODUCTION

STUDIES on fire evacuation seek to understand how occupants react when they are confronted with fire emergencies. Various disciplines, such as safety engineering, computer modeling, human factors, and psychology contribute to this field of research, all aiming to better understand human behavior in fire (HBiF) and ultimately to improve safety. One of the biggest challenges in this field is the access to ecologically valid and at the same time experimentally controlled empirical data (see for example references [1, 2]). Researchers in HBiF need safe, objective, reliable, and valid methods of data collection. The scope of this position paper is to discuss how virtual reality

Paul Pauli and Andreas Mühlberger are shareholders of a commercial company that develops virtual environment research systems for empirical studies in the field of psychology, psychiatry, and psychotherapy. Mathias Müller is executive officer of the same company. No further potential conflicting interests exist.

The presentation of this paper was supported by the German Academic Exchange Service (DAAD).

(VR) studies complement other well established research methods, such as case studies, unannounced drills, field studies, laboratory studies, and hypothetical studies [3-6].

The present article seeks to analyze Strengths, Weaknesses, Opportunities, and Threats (SWOT) of VR research on HBiF. SWOT analysis originates in the management literature and has been applied to VR in the context of rehabilitation research [7].

#### *I. VR in Fire Evacuation*

VR has been defined as a “real or simulated environment in which the perceiver experiences telepresence” (the feeling of being present in a virtual environment) [8]. Note that this very wide definition implies that VR is not limited to computer generated environments or any specific technology. In a way, a real world laboratories can also be seen as virtual environments. However, for the scope of this article, VR refers only to computer generated simulations. The experience of telepresence comprises the illusion of being in the place displayed by the VR technology, and the illusion that events happening in the virtual environment (VE) are plausible and real [9]. Note that this definition does not imply the use of any specific technology. However, VR systems generally use computer generated visual and auditory simulations to immerse participants into a VE. Although less immersive systems – such as simulations on desktop computers [10] – can be used to study HBiF, the present paper mainly addresses highly immersive VR systems using CAVE (Cave Automatic Virtual Environment) systems, Powerwalls, or head mounted displays (HMD). These systems allow the presentation of highly realistic interactive visual and auditory stimuli to participants. Enhanced multimodal systems extend VR with olfactory [11, 12] and proprioceptive stimuli like wind, heat, or motion [13].

VR has become a well-established method in other research fields such as traffic behavior [14], and clinical, social, or experimental psychology, e.g., for the psychotherapy of phobias [15], post-traumatic stress disorder [16], and for rehabilitation [7]. However, the usefulness of VR for HBiF is still under discussion and the method needs to be validated. *Ecological* validity can be assumed, if participants show similar behavioral, emotional, cognitive, and psychophysiological reactions in VR and in the real world [17]. The extent of emotional responses to a real world or virtual laboratory scenario is not necessarily the same as one might expect in a real fire emergency. Ecological validity does not mandate that participants have to believe that a simulated fire scenario is real. In fact, perceptual input (e.g., a visual simulation) can elicit emotional reactions, such as fear reactions, even if participants know that what they see is a simulation [18]. These reactions are probably of a lower intensity, however, future research is necessary to shed light on this question in the context of HBiF. More importantly, VEs have to be designed in a way that the observations from participants' behavior allow valid conclusion for real world scenarios. One study found promising results by comparing participants' behavior in VR evacuation scenario with real world case studies [19]. Other studies found HBiF comparable in conventional laboratory and VR simulated tunnel emergency scenarios [20, 21]. Note that similarity of two forms of artificial experimental methods (VR and classical laboratory studies) does not warrant ecological validity. There are still not enough studies systematically comparing virtual and real HBiF. In related research fields,

however, validation studies have repeatedly demonstrated the ecological validity of VR. For example, several driving simulator studies documented ecological validity of VR simulations in terms of driving behavior, as well as the ability to elicit adequate emotional responses to VR [22-24]. In addition, several validation studies of virtual driving simulators demonstrated similar behavior in the real and the virtual world [25-28].

So far, VR has been used in several studies on diverse aspects of human behavior in fire, such as evacuation from buildings [10, 19, 29-32], occupant behavior in road tunnel fires [33-35], fire training [36-40], and other areas of safety and security research [41-43].

If proven sufficiently valid, VR will be a promising route to gain objective and reliable insights in HBiF. Results from VR studies can be used to test theories of HBiF, verify and validate evacuation models [44], and be integrated into VR training measures [29, 45].

## II. VR in Comparison to other Methods

Table I compares six different empirical research methods, which gather data on HBiF, on several important aspects such as the degree of experimental control, experimental setting, and the type of data that can be collected with each method.

In *hypothetical studies* have been used in evacuation research [46]. Participants are usually either shown videos or are instructed to imagine a certain scenario and then asked how they would react in that situation. Another example would be data acquisition from experts who evaluate the outcome of a given hypothetical scenario. These scenarios can be in the form of an interview or questionnaires. Data

TABLE I.  
COMPARISON OF RESEARCH METHODS

	Hypothetical study	"Classical" lab experiment	VR experiment	Field studies	Drills	Case Studies
<b>Setting</b>	laboratory	laboratory	laboratory	real-world	real-world	real-world
<b>Experimental control</b>	yes	Yes (less than in VR)	yes	limited	no	no
<b>Setting</b>	laboratory	laboratory	laboratory	real-world	real-world	real-world
<b>Type of data</b>	subjective (statements from participants or experts)	subjective, objective (behavior & psycho-physiology)	subjective, objective (behavior & psycho-physiology)	subjective, objective (behavior)	subjective, objective (behavior)	subjective, objective (behavior)
<b>Possibility of use of stressors</b>	no (only hypothetical)	limited	limited	limited	limited	yes
<b>Ecological validity</b>	low	medium	medium	medium	high, if unannounced; limited, if announced	high
<b>Possibility of adjusting experimental setting</b>	yes	yes	yes	limited	no	no
<b>Possibility of exact replication</b>	yes	yes	yes	limited	no	no
<b>Time and cost intensity for data collection</b>	very low	low	low	high	medium	-
<b>Automatic data collection possible</b>	yes	yes	yes	limited	limited	no

from hypothetical studies is always subjective as they reflect the participants' personal opinion, knowledge, or experience. Subjective data, although being prone to bias, can be highly useful to gain insights into how occupants experienced an event or to reconstruct chains of events.

In "*classical*" *laboratory studies* real world scenarios are transferred into the controlled environment of a laboratory. Here, causal effects can be investigated with experimental methods by manipulating independent variables and measuring dependent variables (e.g., behavior, subjective data, and physiological data). Participants have to be assigned randomly into at least two experimental conditions for a true experiment which vary only in one condition (the independent variable). It is crucial that laboratory studies are ethical acceptable. That is, the experimenter may, for example, only use stimuli/stressors that are not actually harming the participant.

In *VR experiments*, participants can be confronted with simulated fire emergencies. Simulations of fire emergencies can be presented to participants in an extremely controlled way. VR experiments allow the convenient recording of behavioral and physiological data with a very high resolution as well as the collection of subjective data. In comparison to other methods, the presentation of stressors (e.g., flames) is ethically less critical compared to classical laboratory and field studies.

In *field studies*, emergency scenarios can be reenacted in a naturalistic setting outside of a laboratory. Unlike in laboratory settings, field studies are usually in less controlled environments (although certain infrastructures like road tunnels are highly controllable). Similar to classical laboratory studies and VR experiments, field studies use subjective and objective data (recorded behavior).

*Drills* are either announced or unannounced practice scenarios in real world settings. Although very similar to field studies, the focus of drills is usually on practicing emergency procedures. They allow the observation of occupant behavior under naturalistic conditions in a specific location. Similar to field studies, observational data and self-report data can be acquired.

*Case studies* refer to the descriptive, exploratory or explanatory analysis of a real fire emergency. Subjective self-report data from occupants and analysis of closed-circuit television footage can be used to reconstruct the events of a real emergency.

In addition, mixed methods may help to overcome limitations of individual methods. For example it is possible to modify participants experience in real world settings using augmented reality, or increase the immersiveness of a VR system by adding real elements (for example objects that participants can touch) to a VR study.

When planning studies on HBiF researchers have to consider certain factors and restrictions (See Table I). These include the necessary degree of experimental control, the choice of setting (laboratory or real-world), or the type of data required (subjective vs objective) and whether or not it is important to be able to adjust or replicate the experimental

scenario during data acquisition. There are also factors related to the efforts necessary for the realization of a study. Efforts can be financial (e.g., costs for hard and software, personnel, participant recruitment, or lab space in VR experiments) but also whether or not data can be collected and processed automatically (e.g., with tracking devices) or has to be extracted from video footage.

These comparisons do not necessarily reflect strengths or weaknesses, rather factors that should be considered when deciding on which research method is most suitable for a certain research question. The methods discussed here do not provide the best solution for every research issue. There are arguments for and against the use of each method to address specific concerns. In this section, the VR studies are analyzed with respect to key aspects of a research question.

## II. SWOT ANALYSIS

SWOT analysis refers to the analysis of **S**trengths, **W**eaknesses, **O**pportunities and **T**hreats of a given method or product [7]. The present SWOT analysis (Table II) aims to uncover internal strengths and weaknesses of VR as a research tool in HBiF and to identify its surrounding conditions (opportunities and threats). A detailed description of SWOT analysis can be found in reference [7].

- Strengths refer to inherent resources and capacities of VR helping to gather objective, reliable, and valid empirical data on HBiF.
- *Weaknesses* describe inherent shortcomings, limitations, and problems of VR to achieve its goal.
- *Opportunities* comprise surrounding conditions or trends from which VR research will potentially benefit and which is promising to overcome weaknesses.
- *Threats* are surrounding conditions which are detrimental to the use of VR as a research tool in HBiF and which need to be overcome.

### I. Strengths

*Internal validity* is possibly the most important strength of VR studies. Entire VEs can be easily controlled. Stimulus control and experimental stimulus manipulation is a key feature in investigating cause and effect relations [47]. It is extremely difficult to impossible to control the environment in field studies, drills, and even classical lab studies. For instance, in VR smoke can be numerically calculated and then repeatedly presented in exactly the same way to several participants. "Real" smoke, even in the controlled environment of a classical laboratory study, will always vary, and consequently visibility conditions may change across observations. Lack of experimental control limits the reliability and consequently the internal validity of these methods.

*Replication.* VR studies can be replicated to the last detail, given the usage of the same or comparable equipment. One major criterion for empirical studies is that they can be/should be reproducible. Replication refers to the

repetition of a study using the same methods but different participants and experimenters. Studies need to be replicated in order to test their reliability and validity and to test their generalizability and the role of confounding variables. Real world studies, especially field and case studies, provide data for only one specific event and are extremely difficult to replicate.

*Ecological validity* refers to the degree with which the methods of a study represent the real world scenario that is being examined. VR offers a similar degree of ecological validity as classical laboratory studies, but depending on the research question one method or the other may be more suitable. For example, certain features of a fire emergency, such as the visual simulation of flames, may be simulated with higher control in VR but other features (e.g. touch) may be more difficult but not impossible to simulate in VR (e.g. using a mix of virtual and real elements). However, simulation of heat or olfactory cues is possible but still limited as it is both technically challenging to present olfactory stimuli in an experimentally controlled manner. Ecological validity of VR studies is higher than in hypothetical studies since the latter require the ability of participants to correctly imagine a scenario. High – but not absolute – ecological validity of VR studies can be assumed if the visualization, observed behavior, and task difficulty of a simulated fire emergency is realistic, i.e., based on valid models and representative of real world events. VR simulations can have the same degree of visual realism as simulations in classical laboratory studies.

All laboratory experiments including VR studies, however, are abstractions of reality and therefore some loss of ecological validity is inherent to the method in comparison to real events [47]. Even the most sophisticated field experiment and the most advanced VR simulation on human behavior in dangerous situations cannot (and should not) claim absolute ecological validity. Participants will always know that they are taking part in an artificial situation. However, this is true for all methods compared in the present article with the exception of unannounced drills and case studies. Knowing that one takes part in a study and/or that there is no real danger, may lead to systematic biases in participants' responses.

*External validity* refers to the question whether the results of a study can be generalized from the experimental setting to other situations or populations [48]. VR and other laboratory studies allow controlling confounding factors and thus studying general underlying effects in HBiF is possible. Results from uncontrolled studies (e.g., drills, case studies, and to some extent also in field studies) cannot not be generalized because confounding variables are not controlled.

*Safety for participants.* VR allows the controlled simulation of perilous scenarios, such as extreme tunnel fires, without putting the participants at risk of a physical harm. That is, VR studies are ethically less problematic than field studies since it is possible to simulate catastrophic and life threatening situations without risking to physical harm

participants. However, there are also limitations for VR studies (see *Threats*).

*Real-time feedback.* Precise tracking of various parameters as well as the highly controlled visual input technologies allow to give participants and researchers immediate feedback of behavior, performance, and even psychophysiological processes. For example, task-performance or physiological parameters such as heart rate can be displayed online during trials. This allows the experimenter to have real time access to data. Real time feedback for participants can be used to test training measures (e.g., fire evacuation training).

*Multi-modal simulations.* In theory, simulation of any modality is possible. To date, combined simulation of visual and auditory stimuli are very well developed. Olfactory, nociceptive, or thermoceptive simulations are also possible, however, still less technologically advanced.

*Precise measurement.* Precise tracking technology allows accurate analysis of various aspects of participants' behavior (e.g., full body tracking, head movement, eye tracking) with extremely high sampling rates.

*Psychophysiological monitoring.* In addition to behaviors, psychophysiological parameters such as heart rate or skin conductance can be measured easily in a VR laboratory. Measuring physiological correlates of behavior while being immersed into a VE allows researchers to analyze emotional reactions (e.g., fear reactions) to simulated emergencies.

*Low costs.* Once a VR system is set up, it can be used, in theory, infinitely. Virtual scenarios can be re-used and easily modified. With the decrease in prices for hardware and software (some VR simulation software are even free to use), VR experiments have become more and more affordable. Although costs for individual studies vary significantly, VR studies are generally cheaper than field studies. However, setting up a complete complex VR laboratory such as CAVE systems is cost intensive and requires space but is relatively affordable to run.

*Repeated measurements.* Participants can easily be immersed repeatedly into VEs and repeated measurements in identical scenarios are possible. Recreating identical conditions is complex or even impossible with other methods (See also *Replication*). Repeated measurements can be used to test, for example, training measures aimed to improve HBiF.

*Flexibility.* Experimental settings in VR can be adjusted easily, allowing to run pilot studies and to quickly develop minor alterations of the experimental set-up.

*Control of confounding variables.* There are many variables that potentially confound the effect of a given independent variable but are not of primary interest (e.g., minor changes in starting positions, left/right turning preferences). These can be easily controlled in VR and laboratory studies but is difficult to impossible in other designs.

*Independent of imagination abilities/willingness of participants.* Producing highly immersive VEs reduces the variance in participants' response caused by individual

differences in the ability and willingness to imagine a given scenario. Hypothetical studies rely on the ability of participants to imagine a scenario. Here, researchers have no control of the ability and willingness of participants to imagine, for example, a fire evacuation from a high-rise building.

*Participant recruitment.* Although not an important strength of VR studies, it is worth mentioning that recruiting a sample for a VR study has less restrictions than recruiting for a drill or field study. In field studies, the experimental set-up is often only available on limited occasions and may be time and cost intensive to install. For example, certain infrastructures such as underground public transportation systems or road tunnels may only be accessible to researchers at a very limited time or during certain hours of the day making data collection difficult. Once a VR scenario has been set-up it can be repeated, in theory, at any given time which allows longer and more flexible time-windows for data collection.

**II. Weaknesses**

*Need for confirmation/validation.* To date, there are still not enough validation studies. These are necessary to test the assumptions that behavior in a simulated scenario can predict or be transferred to “real” HBiF.

*Non-intuitive interaction methods.* Although VEs are interactive, participants often need devices like gamepads to navigate in and interact with the virtual environment. This always reminds the participants that they are in an artificial scenario, and if the scenario is not well designed, may bias behavior. For example, evacuation times may vary depending on how well participants can handle their navigation device. However, developments in input devices may partially address this weakness (See also *Opportunities*).

*Inter-individual differences in ease of interaction with VR.* Depending on various factors, such as age or experience with VR, participants may have difficulties when using VR. For example, participants who have a lot of experience using 3D video games may find it easier to navigate in a VE. Participants, who have less experience with computers (e.g., elderly participants) may need longer practice sessions before they can navigate without limitations in a VE.

*Technical limitations.* Visual input as well as interaction methods are still limited. Although visual simulation of virtual environments has improved tremendously in recent years, the current simulations will always be recognized as such by participants. Such imperfections (e.g. in model rendering, spatial resolution, field of view (for HMDs), graphic update rate, lags between head tracking and visualization) of VEs may lead to artifacts [49]. Especially the simulation of behaviorally realistic virtual humans is still challenging. Another technological limitation is the need for interaction tools, such as game pads or HMDs, to immerse into and interact in the VE. For example, navigation, even in a highly immersive CAVE system is either limited to a few square meters or participants have to use interaction devices. These technical challenges may limit the immersiveness of a VR system and lead to a lower experience of presence.

*Technology-induced side effects.* Prolonged exposure to VR may cause symptoms of nausea and vertigo (simulator sickness; for a review on simulator sickness, see references [49, 50]). The incidence of these side effects depends upon characteristics of the VR system (e.g., display field of view, lag between tracking device and update of the visualization) and participants [51]. Such side-effects need to be considered when planning and evaluating the ethical innocuousness (e.g., participants need to be able to terminate the experiment whenever they want to).

TABLE II.  
SUMMARY OF A SWOT ANALYSIS FOR VR IN FIRE EVACUATION RESEARCH.

Strengths	Weaknesses	Opportunities	Threats
<ul style="list-style-type: none"> <li>• Internal validity</li> <li>• Replication</li> <li>• Ecological validity</li> <li>• External validity</li> <li>• Safety for participants</li> <li>• Real-time feedback</li> <li>• Multi-modal simulations</li> <li>• Precise measurement</li> <li>• Psychophysiological monitoring</li> <li>• Low costs</li> <li>• Repeated measurements</li> <li>• Flexibility</li> <li>• Control of confounding variables</li> <li>• Independent of imagination abilities/willingness of participants</li> <li>• Participant recruitment</li> </ul>	<ul style="list-style-type: none"> <li>• Need for confirmation/validation</li> <li>• Non-intuitive interaction methods</li> <li>• Inter-individual differences in ease of interaction with VR</li> <li>• Technical limitations</li> <li>• Technology-induced side effects</li> <li>• Efforts</li> </ul>	<ul style="list-style-type: none"> <li>• Intuitive and natural navigation</li> <li>• Graphical developments</li> <li>• Multi-modal simulation and feedback</li> <li>• Usability for researchers</li> <li>• Exchange of 3D-scenes or experiments</li> </ul>	<ul style="list-style-type: none"> <li>• Failure to show ecological validity</li> <li>• Ethical challenges</li> <li>• Side-effects due to interaction with other medical conditions</li> <li>• Misleading expectations</li> <li>• Technical faults</li> </ul>

*Efforts.* Setting up a highly immersive VR laboratory is time and cost intensive. [49]. Creating plausible VEs is also complex (the differences between less immersive and simple environments and complex highly immersive VEs is extreme) and requires expertise with special hard- and software systems. Given the rapid developments in this type of technology, constant investment may be required to stay up-to-date.

### III. Opportunities

*Intuitive and natural navigation.* Although, highly immersive VR systems, such as CAVE or HMD systems, allow participants to move freely with their whole body within the VE [52], even the most advanced systems still have movement restrictions and participants have to use navigation and input devices. Advances in tracking technology, innovative interaction devices, and VR systems that allow natural navigation (e.g., bigger CAVE systems or wireless HMDs and walking platforms) may reduce the weakness of limited space and non-intuitive interaction devices, e.g. [52]. These advances promise even better immersion into VEs and consequently improved ecological validity.

*Graphical developments.* The dramatic improvement of graphical simulations allows more and more photorealistic simulation of fire emergencies. In addition, numerical calculated fire and smoke have been successfully implemented into VR simulations [37]. Similar to the advances in navigation devices, improved realism of simulations will lead to increased experience of presence for participants and better ecological validity.

*Multi-modal simulation and feedback.* The integration of multi-modal simulations for visual and auditory simulation extended by kinesthetic, olfactory, haptic, thermoceptive simulation allows the simulation of more complete scenarios. For examples, see references [53, 54].

*Usability for researchers.* The widespread use of VR technology depends highly on its usability for researchers. Recent developments in easy to use VR tool kits make VR technology more accessible. The improved cross platform compatibility helps the use of VR over different platforms and operation systems.

*Exchange of 3D-scenes or experiments.* Researchers can easily exchange 3D models or even entire experiments with each other. This may foster cooperation between laboratories and also lead to the development of standard scenarios which could be used as references and thus increase comparability of different VR studies.

### IV. Threats

*Failure to show ecological validity.* This is the biggest threat to VR as a research tool to study HBiF. Systematic validation of VR for HBiF has still not demonstrated its range of applicability. Future studies are clearly necessary to test the ecological validity of VR to study HBiF.

*Ethical challenges.* Scientific studies on HBiF have to comply with ethical standards such as the Declaration of Helsinki which define ethical standards for studies with

human subjects [55]. Even though most participants are aware that a virtual fire provides no threat to them, some participants may still experience extreme fear. Just as with any other method, VR research needs to ensure that the experienced fear cannot lead to longer lasting difficulties for participants such as traumatization, especially if one has in mind that VEs are getting closer and closer in means of realism to real scenarios. In addition, a VR system that causes extreme side effects (e.g., seizures or strong nausea) would be ethically unacceptable.

If participants cannot differentiate between a simulated and a real scenario, which may be the case, for example, with small children, the same ethical restrictions as with other methods apply.

*Side-effects due to interaction with other medical conditions.* Some scenario for HBiF may be particularly risky in causing side-effects in interaction with pre-existing medical conditions. For example, studies using flashing lights may cause seizures in at-risk populations; patients with specific phobias (e.g., of tunnels or heights) may experience extreme fear; Simulation of fire emergencies may induce flashbacks in participants who previously have experienced a traumatizing event. Other methods, however, bear similar risks.

*Misleading expectations.* The expectation that VR experiments can completely replace real world tests and holistically covers all aspects of human behavior in fire is misleading. Similar to classical laboratory experiment, VR allows investigating general underlying processes of HBiF and testing specific aspects (e.g., the effect of safety installations on evacuation behavior). The conclusions from these studies may even lead to changes in the design of real world safety installations. However, HBiF is highly complex; one can never exclude that individual decision-making, behavior, and experience in a specific real scenario may differ significantly from trends found in VR studies.

Certain *technical faults* in the implementation of a VR system (e.g., jitter errors, discrepancies in simulation or tracking latency) even can reduce the immersiveness and even may increase side-effects like simulator sickness.

## III. GENERAL DISCUSSION

The present position paper provides a SWOT analysis for VR as a research tool to study HBiF. We provided an overview of various methods used in HBiF and systematically compared VR to these methods.

The biggest strength of VR is surely its ability to create highly immersive, externally valid, highly controlled, and safe experimental set-ups. The biggest weakness is the reduced ecological validity in comparison with field and case studies, as well as the lack of validation studies specifically for HBiF. These studies should compare VR experiments with the results of other laboratory experiments and field studies.

The diverse methods used to study HBiF always have to trade-off between ecological validity and experimental control. For instance, case and field studies in real world

settings provide almost perfect ecological validity. However, strict experimental control is impossible to achieve here, and financial and logistic efforts as well as ethical limitations need to be considered. Hypothetical studies need to consider less strict ethical limitations and are easier to realize, but rely heavily on the ability of the participant's imagination and are prone to response biases.

Field studies are often characterized by the combination of setting and participants, e.g., real world settings with participants that naturally are in these settings. In the evacuation area, this allows doing unannounced experiments. This can never be achieved in VR or other laboratory experiments, as participants need to be recruited and enter the VR-lab (or ask them to put on some equipment). At most, participants may be "deceived" by telling them that they will take part in one study and then exposing them to something else. However, this is easily feasible in classical laboratory studies but requires more efforts in VR studies. It can be argued that affects the external validity of a study.

The differentiation between ecological and external validity is important. Ecological validity refers to how good a research method represents reality. External validity describes how well study results can be transferred to other situations and generalized over populations. Whereas ecological validity is not, external validity is a prerequisite for the overall validity of a study. A study, can be ecologically valid (e.g., the results from an unannounced drill) but not generalizable to other settings, populations, cultures etc., if it lacks experimental control and, therefore, internal validity).

### I. What can we study in VR?

VR can be used to design complex laboratory experiments on HBiF. It allows studying how occupants react to fire cues, such as flames or smoke; it allows collecting precise behavioral and psychophysiological data during controlled simulated events. Virtual scenarios can be designed with an extremely high level of detail. That way, we can use VR to study underlying processes of HBiF (e.g. phenomena like risk perception of occupants, social influence, architectural influences, way-finding abilities in smoke, etc.). That way, VR studies can contribute to a better understanding of HBiF.

In addition, evacuation concepts for large complex buildings can be tested in VR making it possible to identify potentially problematic evacuation routes *before* a new building is constructed. This is particularly useful since evacuation models implemented in simulation software tools still oversimplify HBiF (e.g., some models assume that occupants always take the shortest route to an emergency exit [33]).

It is important to note that VR cannot replace any of the other methods mentioned above but is complementary. VR studies can be used in experimental pilot studies in order to test a number of possible factors that may theoretically be influencing HBiF (e.g. various design aspects of safety equipment). Then, those factors deemed as the most

important ones in VR can then be tested in field experiments or used to predict behavior in case studies or drills.

### II. What can we not study in VR?

Virtual reality is not reality. Participants will always know that they take part in an artificial situation. It is impossible to generate situations in which participants' would risk actual physical harm. Extremely perilous situations may induce effects (e.g., extreme fear) which are not attainable with artificial scenarios, which in turn may affect behavior. Only observations from real events and to some degree unannounced drills may have this effect. It is impossible to investigate these parts of HBiF using VR laboratory studies.

### III. Conclusion and positioning statement

We argue that VR is a powerful approach to study HBiF. VR allows shedding light on aspects of occupant behavior that were previously impossible to investigate under controlled conditions. Although we identified several weaknesses and limitations of the method, the most important one being the need for validation studies, it seems possible that these can be overcome, either by technical progress or by combining several different research approaches (triangulation approach). None of the state of the art research methods (including VR) are able to validly grasp all aspects of HBiF, and VR does not aim to replace any of the other presently established research methods. We see it as a promising complementary laboratory tool in the quest to understand HBiF and to improve fire safety.

### IV. REFERENCES

- [1] K. Fridolf, D. Nilsson, and H. Frantzich, "Fire evacuation in underground transportation systems: a review of accidents and empirical research," *Fire Technology*, vol. 49, pp. 451-475, 2013.
- [2] M. Kobes, I. Helsloot, B. de Vries, and J. G. Post, "Building safety and human behaviour in fire: A literature review," *Fire Safety Journal*, vol. 45, pp. 1-11, 1// 2010.
- [3] R. F. Fahy and G. Proulx, "A comparison of the 1993 and 2001 evacuations of the World Trade Center," in *Proceedings of the 2002 Fire Risk and Hazard Assessment Symposium*, 2002, pp. 111-117.
- [4] D. Nilsson, H. Frantzich, and W. Saunders, "Coloured Flashing Lights to Mark Emergency Exits - Experiences from Evacuation Experiments," presented at the Fire Safety Science - Proceedings of the Eighth International Symposium, Beijing, China, 2005.
- [5] T. Shields and K. Boyce, "A study of evacuation from large retail stores," *Fire Safety Journal*, vol. 35, pp. 25-49, 2000.
- [6] P. Burns, G. Stevens, K. Sandy, A. Dix, B. Raphael, and B. Allen, "Human behaviour during an evacuation scenario in the Sydney Harbour Tunnel," *Australian Journal of Emergency Management, The*, vol. 28, p. 20, 2013.
- [7] A. S. Rizzo and G. J. Kim, "A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy," *Presence: Teleoperators and Virtual Environments*, vol. 14, pp. 119-146, 2005/04/01 2005.
- [8] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *Journal of communication*, vol. 42, pp. 73-93, 1992.
- [9] M. Slater, B. Spanlang, and D. Corominas, "Simulating virtual environments within virtual environments as the basis for a psychophysics of presence," *ACM Transactions on Graphics (TOG)*, vol. 29, p. 92, 2010.
- [10] N. W. Bode, A. U. K. Wagoum, and E. A. Codling, "Human responses to multiple sources of directional information in

- virtual crowd evacuations," *Journal of The Royal Society Interface*, vol. 11, p. 20130904, 2014.
- [11] W. Barfield and E. Danas, "Comments on the use of olfactory displays for virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 5, pp. 109-121, 1996.
- [12] E. Richard, A. Tijou, P. Richard, and J.-L. Ferrier, "Multi-modal virtual environments for education with haptic and olfactory feedback," *Virtual Reality*, vol. 10, pp. 207-225, 2006.
- [13] F. Hülsmann, N. Mattar, J. Fröhlich, and I. Wachsmuth, "Wind and Warmth in Virtual Reality—Requirements and Chances," in *Proceedings of the Workshop Virtuelle & Erweiterte Realität 2013*, 2013.
- [14] L. N. Boyle and J. D. Lee, "Using driving simulators to assess driving safety," *Accident Analysis and Prevention*, vol. 42, pp. 785-787, May 2010.
- [15] K. Meyerbröker and P. M. Emmelkamp, "Virtual reality exposure therapy in anxiety disorders: a systematic review of process-and-outcome studies," *Depression and anxiety*, vol. 27, pp. 933-944, 2010.
- [16] B. K. Wiederhold and M. D. Wiederhold, "Virtual Reality Treatment of Posttraumatic Stress Disorder Due to Motor Vehicle Accident," *Cyberpsychology Behavior and Social Networking*, vol. 13, pp. 21-27, Feb 2010.
- [17] C. A. Anderson and B. J. Bushman, "External validity of "trivial" experiments: The case of laboratory aggression," *Review of General Psychology*, vol. 1, pp. 19-41, 1997.
- [18] H. M. Peperkorn, G. W. Alpers, and A. Mühlberger, "Triggers of Fear: Perceptual Cues Versus Conceptual Information in Spider Phobia," *Journal of clinical psychology*, 2013.
- [19] M. Kobes, I. Helsloot, B. de Vries, and J. Post, "Exit choice, (pre-)movement time and (pre-)evacuation behaviour in hotel fire evacuation — Behavioural analysis and validation of the use of serious gaming in experimental research," *Procedia Engineering*, vol. 3, pp. 37-51, 2010/01// 2010.
- [20] F. Malthe and Vukancic, "Virtual Reality och människors beteende vid brand [Virtual Reality and human behavior in fire]," Lund University LUCATORG: 011033007, 2012.
- [21] J. Johansson and L. Petersson, "Utrymning och vägval i Virtual Reality."
- [22] A. Mühlberger, H. H. Bühlhoff, G. Wiedemann, and P. Pauli, "Virtual reality for the psychophysiological assessment of phobic fear: responses during virtual tunnel driving," *Psychological Assessment*, vol. 19, pp. 340-346, Sep 2007.
- [23] A. Calvi and M. R. De Blasiis, "How Long is Really a Road Tunnel? Application of Driving Simulator for the Evaluation of the Effects of Highway Tunnel on Driving Performance," in *6th International Conference Traffic and Safety in Road Tunnels*, Hamburg, Germany, 2011.
- [24] A. Calvi, "Analysis of Driver's Behaviour in Road Tunnels: a Driving Simulation Study," in *2010 International Symposium on Safety Science and Technology*, Zhejiang, China, 2010.
- [25] J. Törnros, "Driving behaviour in a real and a simulated road tunnel - A validation study," *Accident Analysis and Prevention*, vol. 30, pp. 497-503, Jul 1998.
- [26] T. Hirata, T. Yai, and T. Tagakawa, "Development of the driving simulation system MOVIC-T4 and its validation using field driving data," *Tsinghua Science & Technology* vol. 12, pp. 141-150, 2007.
- [27] O. Shechtman, S. Classen, K. Awadzi, and W. Mann, "Comparison of Driving Errors Between On-the-Road and Simulated Driving Assessment: A Validation Study," *Traffic Injury Prevention*, vol. 10, pp. 379-385, 2009.
- [28] S. Heliovaara, J.-M. Kuusinen, T. Rinne, T. Korhonen, and H. Ehtamo, "Pedestrian behavior and exit selection in evacuation of a corridor - An experimental study," *Safety Science*, vol. 50, pp. 221-227, Feb 2012.
- [29] U. Rüppel and K. Schatz, "Designing a BIM-based serious game for fire safety evacuation simulations," *Advanced Engineering Informatics*, vol. 25, pp. 600-611, 2011.
- [30] E. Duarte, F. Rebelo, J. Teles, and M. S. Wogalter, "Behavioral compliance for dynamic versus static signs in an immersive virtual environment," *Applied Ergonomics*, in press.
- [31] G. Lawson, S. Sharples, D. Clarke, and S. Cobb, "Validating a low cost approach for predicting human responses to emergency situations," *Applied Ergonomics*, vol. 44, pp. 27-34, 1// 2013.
- [32] L. Gamberini, P. Cottone, A. Spagnolli, D. Varotto, and G. Mantovani, "Responding to a fire emergency in a virtual environment: different patterns of action for different situations," *Ergonomics*, vol. 46, pp. 842-858, Jun 20 2003.
- [33] E. Ronchi, M. Kinateder, M. Müller, M. Jost, M. Nehfischer, P. Pauli, *et al.*, "Evacuation travel paths in virtual reality experiments for tunnel safety analysis," submitted.
- [34] M. Kinateder, E. Ronchi, M. Müller, M. Jost, M. Nehfischer, P. Pauli, *et al.*, "Social influence on route choice in a virtual reality tunnel fire," submitted.
- [35] M. Kinateder, M. Müller, A. Mühlberger, and P. Pauli, "Social Influence in a Virtual Tunnel Fire - Influence of Passive Virtual Bystanders," in *Human Behaviour in Fire 2012*, Cambridge, 2012, pp. 506-516.
- [36] M. Kinateder, P. Pauli, M. Müller, J. Krieger, F. Heimbecher, I. Rönna, *et al.*, "Human behaviour in severe tunnel accidents: Effects of information and behavioural training," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 17, pp. 20-32, 2013.
- [37] Z. Xu, X. Z. Lu, H. Guan, C. Chen, and A. Z. Ren, "A virtual reality based fire training simulator with smoke hazard assessment capacity," *Advances in Engineering Software*, vol. 68, pp. 1-8, 2// 2014.
- [38] K. M. O'Connell, M. J. De Jong, K. M. Dufour, T. L. Millwater, S. F. Dukes, and C. L. Winik, "An Integrated Review of Simulation Use in Aeromedical Evacuation Training," *Clinical Simulation in Nursing*, vol. 10, pp. e11-e18, 1// 2014.
- [39] S. L. Farra, E. T. Miller, and E. Hodgson, "Virtual reality disaster training: Translation to practice," *Nurse Education in Practice*.
- [40] M. Cha, S. Han, J. Lee, and B. Choi, "A virtual reality based fire training simulator integrated with fire dynamics data," *Fire Safety Journal*, vol. 50, pp. 12-24, 5// 2012.
- [41] A. I. Ginnis, K. V. Kostas, C. G. Politis, and P. D. Kaklis, "VELO: A VR platform for ship-evacuation analysis," *Computer-Aided Design*, vol. 42, pp. 1045-1058, 11// 2010.
- [42] K. Andree, M. Kinateder, and D. Nilsson, "Immersive Virtual Environment as a Method to experimentally study human behaviour in fire," in *13th International Conference and Exhibition on Fire Science and Engineering*, Royal Holloway College, University of London, UK, 2013, pp. 565-570.
- [43] J. Drury, C. Cocking, S. Reicher, A. Burton, D. Schofield, A. Hardwick, *et al.*, "Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model," *Behavior research methods*, vol. 41, pp. 957-970, 2009.
- [44] T. Kretz, S. Hengst, A. P. Arias, S. Friedberger, and U. D. Hanebeck, "Using a Telepresence System to Investigate Route Choice Behavior," *arXiv preprint arXiv:1111.1103*, 2011.
- [45] J. Ribeiro, J. E. Almeida, R. J. Rossetti, A. Coelho, and A. L. Coelho, "Using Serious Games to Train Evacuation Behaviour." W. Saunders, "Decision making model of behaviour in office building fire evacuations," PhD thesis, Department of Psychology, Victoria University of Technology, 2001.
- [47] J. M. Loomis, J. J. Blascovich, and A. C. Beall, "Immersive virtual environment technology as a basic research tool in psychology," *Behavior Research Methods, Instruments, & Computers*, vol. 31, pp. 557-564, 1999.
- [48] W. R. Shadish, T. D. Cook, and D. T. Campbell, "Experimental and quasi-experimental designs for generalized causal inference," 2002.
- [49] J. Loomis, J. Blascovich, and A. Beall, "Immersive virtual environment technology as a basic research tool in psychology," *Behavior Research Methods*, vol. 31, pp. 557-564, 1999.
- [50] R. Patterson, M. D. Winterbottom, and B. J. Pierce, "Perceptual issues in the use of head-mounted visual displays," *Human Factors*, vol. 48, pp. 555-573, Fal 2006.
- [51] K. M. Stanney and R. S. Kennedy, "Simulation Sickness," in *Human factors in simulation and training*, P. A. Hancock, D. A. Vincenzi, J. A. Wise, and M. Mouloua, Eds., ed Boca Raton, Florida: CRC Press, 2010, pp. 117-124.



- [52] A. Nybakke, R. Ramakrishnan, and V. Interrante, "From virtual to actual mobility: Assessing the benefits of active locomotion through an immersive virtual environment using a motorized wheelchair," in *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, 2012, pp. 27-30.
- [53] B. Weber, M. Sagardia, T. Hulin, and C. Preusche, "Visual, Vibrotactile, and Force Feedback of Collisions in Virtual Environments: Effects on Performance, Mental Workload and Spatial Orientation," in *Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments*. vol. 8021, R. Shumaker, Ed., ed: Springer Berlin Heidelberg, 2013, pp. 241-250.
- [54] Heidelberg, 2013, pp. 241-250. J. Hummel, J. Dodiya, R. Wolff, A. Gerndt, and T. Kuhlen, "An evaluation of two simple methods for representing heaviness in immersive virtual environments," in *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, 2013, pp. 87-94.
- [55] W. M. Association, "World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects," *Bulletin of the World Health Organization*, vol. 79, p. 373, 2001.